

38

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-224520

(43) 公開日 平成10年(1998) 8月21日

(51) Int. CL ⁶	識別記号	P I	
H 0 4 M 17/00		H 0 4 M 17/00	Z
G 1 0 L 3/00		G 1 0 L 3/00	R
H 0 4 M 11/06		H 0 4 M 11/06	

審査請求 未請求 請求項の数 2 O L (全 10 頁)

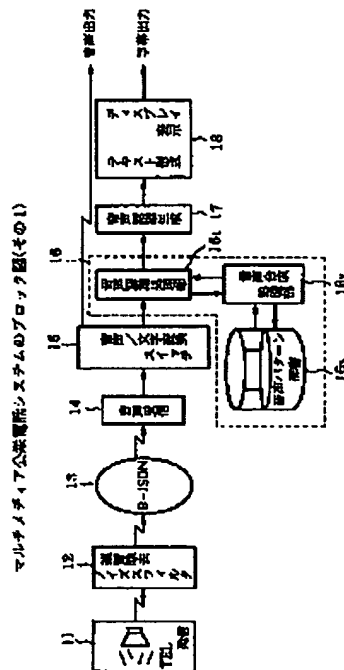
(21) 出願番号	特願平9-26218	(71) 出願人	000006105 株式会社明電舎 東京都品川区大崎2丁目1番17号
(22) 出願日	平成9年(1997) 2月10日	(72) 発明者	高橋 和彦 東京都品川区大崎2丁目1番17号 株式会社明電舎内
		(74) 代理人	弁理士 志賀 富士弥 (外1名)

(54) 【発明の名称】 マルチメディア公衆電話システム

(57) 【要約】

【課題】 マルチメディア公衆電話システムは、映像も送受信できるが、難聴者の場合、会話を聞く事ができないなど、聴覚障害者にとって利用する範囲が限られている。

【解決手段】 通話者の音声を通信システムを介して音声信号受信手段11~14で受信し、音声/文字変換スイッチ15で字幕受信を選択したとき、音声認識装置16により受信した音声信号を認識し、この認識音声を生音声認識生成装置17で文字データに変換してディスプレイ上に字幕で表示する。逆に、通話者が入力した文字列を文字認識により文体として認識生成し、この文章を音声合成により音声信号に変換し、通信システムを介して通話相手に送信する手段も含むことができるようにする。



【特許請求の範囲】

【請求項1】 音声と映像の入出力手段と通信手段を有して互いに音声及び映像を使った通話を可能にしたマルチメディア公衆電話システムにおいて、

通話者の音声を通信システムを介して受信する音声信号受信手段と、

前記音声信号をそのまま音声で出力又は字幕で出力するかを選択する音声／文字変換スイッチと、

前記スイッチが字幕出力を選択したとき、受信した音声信号を認識する音声認識装置と、

前記音声認識装置が認識した音声文字データに変換して前記ディスプレイ上に字幕で表示する音声認識生成装置とを備えたことを特徴とするマルチメディア公衆電話システム。

【請求項2】 音声と映像の入出力手段と通信手段を有して互いに音声及び映像を使った通話を可能にしたマルチメディア公衆電話システムにおいて、

通話者の音声信号をそのまま通話相手に送信するか又は文字で送信するかを選択する文字／音声変換スイッチと、

前記スイッチが文字送信を選択したとき、通話者が入力した文字列を文字認識により文法として認識生成する文章理解装置と、

前記文章理解装置が生成した文章を音声合成により音声波形に変換する音声合成装置と、

前記音声合成装置が合成した音声波形を音声信号に変換して通信システムを介して通話相手に送信する音声変換装置とを備えたことを特徴とするマルチメディア公衆電話システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、公衆電話システムにおいて、音声の他に映像等を送受信できるようにしたマルチメディア公衆電話システムに関する。

【0002】

【従来の技術】 現在のマルチメディア公衆電話システムは、図12に示すように、CCDカメラ1による映像入力とスクリーン2による映像出力になる映像入出力装置により、相手の顔を見ながら同時に会話ができると、FAX機能を持つこと、設置場所の周辺マップや情報案内機能を持つこと、メモ機能を持つこと等の機能が加えられ、さらに天気予報が聞けること、各イベントの情報を取り寄せること等の各種の機能がある。この公衆電話の操作入力は、プッシュボタン3とポインティングデバイス（ペン入力やタッチスクリーン方式）4が利用される。

【0003】

【発明が解決しようとする課題】 従来のマルチメディア公衆電話システムでは、以下の課題がある。

【0004】 (1) 通話者の顔を見ることができても、聴覚者の場合、会話を聞く事ができない。

【0005】 (2) 通話者の音声（会話の内容）を文字として読み取ることができない。

【0006】 (3) 言語障害があっても会話が難しい聴覚者の場合、メモ機能などで話の内容を伝えるために文字で書き表した内容を音声で相手に伝える事ができない。

【0007】 (4) 健常者でも会話中、聞き間違えたり、聞き取れなかったりすることがあり、話を中断する事もある。

【0008】 以上のように、従来のシステムは種々のマルチメディア機能があっても聴覚障害者にとって利用できる範囲が限られている。

【0009】 本発明の目的は、聴覚障害者等にも使い勝手を良くしたマルチメディア公衆電話システムを提供することにある。

【0010】

【課題を解決するための手段】 本発明は、上記課題を解決するため、従来の機能に音声理解システムを追加し、相手の音声信号を字幕に変換処理してディスプレイ上に表示させ、さらに言語系のハンディも緩和させるためにメモ機能に音声理解機能を追加することでメモに記述した文字を音声に変換して通話者に伝達できるようにしたもので、以下の構成を特徴とする。

【0011】 (第1の発明) 音声と映像の入出力手段と通信手段を有して互いに音声及び映像を使った通話を可能にしたマルチメディア公衆電話システムにおいて、通話者の音声信号を通信システムを介して受信する音声信号受信手段と、前記音声信号をそのまま音声で出力又は字幕で出力するかを選択する音声／文字変換スイッチと、前記スイッチが字幕出力を選択したとき、受信した音声信号を認識する音声認識装置と、前記音声認識装置が認識した音声文字データに変換して前記ディスプレイ上に字幕で表示する音声認識生成装置とを備えたことを特徴とする。

【0012】 (第2の発明) 音声と映像の入出力手段と通信手段を有して互いに音声及び映像を使った通話を可能にしたマルチメディア公衆電話システムにおいて、通話者の音声信号をそのまま通話相手に送信するか又は文字で送信するかを選択する文字／音声変換スイッチと、前記スイッチが文字送信を選択したとき、通話者が入力した文字列を文字認識により文法として認識生成する文章理解装置と、前記文章理解装置が生成した文章を音声合成により音声波形に変換する音声合成装置と、前記音声合成装置が合成した音声波形を音声信号に変換して通信システムを介して通話相手に送信する音声変換装置とを備えたことを特徴とする。

【0013】

【発明の実施の形態】 図1及び図2は、本発明の実施形態を示すマルチメディア公衆電話システムのブロック図であり、図3～図6に各部の処理過程を同じ符号で対応付けて示す。以下、各部の構成と処理を詳細に説明す

る。

【0014】(1)相手の音声を手書き表示で受信する場合(図1)。

【0015】音声入力装置11は、公衆電話に搭載される受話器のマイクロフォンから音声を入力するためのものであり、通話者が受話器から音声を入力した信号を発生する。雑音除去フィルタ12は、会話の中に紛れて入ってくる雑音をノイズフィルタで除去する。

【0016】通信システム13は、例えば、デジタル回線で高速かつ広帯域性がある2B+DのB-ISDN及びネットワークを制御するATM交換システムによって通信制御を行い、音声入力装置11からの音声信号を音声受信装置14に伝送する。

【0017】音声/文字変換スイッチ15は、受信した音声(アナログ信号)をそのまま音声として受信するか、文字(デジタル信号)に変換してディスプレイに表示させるかを受信者が選択する。

【0018】音声理解装置16は、スイッチ15の選択操作が文字変換にされた場合に受信音声を認識及び理解する。この音声理解装置16は、受信した音声を認識及び理解するため、不特定話者に対応可能な技術として音声分析とファジィ制御によるあいまいな言語を除去する特徴抽出後に単語・構文・意味解析による音声認識を行うシステムに構成され、音声認識処理部16、と音声合成処理部16、及び音声パターン辞書16、を備える。

【0019】この装置16では、図3及び図4の処理フローで示すように、入力された音声信号をA/D変換処理した後、音声分析により音声入力信号を周波数スペクトルによって大まかに分析し、特徴抽出によって音声分析結果から音声信号の特徴パラメータを時系列に抽出・変換し、セグメンテーション処理により音声単位へのセグメント化を行う。

【0020】特徴抽出には、「えーと」や「あー」、「うーん」のように、冗長的な単語をファジィ推論により抽出して除去し、人と人が会話をする上で本当に必要な単語のみを抽出する。さらに、日常的に必要な単語は前もって登録しておき、後に必要となった単語を学習機能によってしだいに認識及び追加して行く。

【0021】次いで、音声認識により音声標準パターンとの比較で単語認識を行うことで音素系列を得、単語照合と単語認識により音素系列について知識ベースに持つ単語標準パターンとの照合で単語を認識する。この認識に標準パターンが存在しなければ学習処理(認識・登録)を行って単語知識を知識ベースに追加する。

【0022】次いで、認識した単語について、構文照合と構文認識により知識ベースの構文パターンとの照合で構文的に誤りがあるか否かを解析し、誤りがあれば再検証することで構文認識を行う。さらに、意味解析と意味認識により認識された単語構文について意味的に妥当であるかを調べ、妥当な結果が得られるまで構文解析処理

と意味解析処理を繰り返す。これら解析には文字に変換可能か否かも含め、文字変換処理により変換可能である文字を漢字やカナに変換する。

【0023】図1に戻って、音声認識生成装置17は、音声理解装置16によって受信音声の内容が認識された音声データをテキスト形式の文字データに変換する。この文字データは、ディスプレイ表示装置18により公衆電話のスクリーン2に字幕で画面表示される。

【0024】したがって、図1の構成により、受信した音声を手書き表示装置により音声認識を行い、これを文字としてスクリーンに表示することにより、難聴者の場合でも通話者の会話を文字として読み取ることができる。

【0025】また、健康者でも電話でのコミュニケーション中、字幕表示により聞き漏らすなどのリスク回避を可能にする。

【0026】(2)文字入力を相手に音声で送信する場合(図2)。

【0027】文字/音声変換スイッチ19は、会話の受信者が、送信者に返答するために、メモ機能で文字で書き表した内容でコミュニケーションを行うときにその操作で文字ボタンを選択する。

【0028】文字入力装置20は、スイッチ19の文字ボタンを押すと、メモ帳(テキスト形式)を画面表示し、話の内容をポインティングデバイス(電子ペン)で入力可能にする。

【0029】文章理解装置21は、手書き入力した文字列を解析して文章として認識するもので、文体認識処理部21、と文体形式分析部21、と文字パターン認識辞書21、及び文体認識生成部21、を備える。

【0030】この文章理解装置21では、図5に処理フローを示すように、手書き文字入力に対して文字認識処理部21、が文字パターンと文体形式から文字の解析を行い、分析内容と入力内容が違えば訂正として再入力を受け、手書き文字の入力内容を文章として認識し、文体認識生成部21、により文章として生成する。

【0031】音声合成装置22は、文章理解装置21で認識生成された文章を音声合成により音声信号に変換するもので、音声合成システム部22、と音声パターン辞書22、を備える。

【0032】音声合成装置22は、規則合成方式により音声合成を行い、漢字カナ混じり文に対して日本語辞書を参照した構文解析、意味解析等により読み・単語・文節境界等を解析し、この解析結果から音声パターン辞書を参照して音素系列のアクセントとイントネーション及び音韻の継続時間の各パラメータを決定し、これらを音響モデルのパラメータとして音声合成を行う。

【0033】音声変換部23は、音声信号波形を音声に変換する。この変換において、通話者本人が男性の声か女性の声かをシステム上で選択し送信ボタンを押す、台成した音声を通信用システム24を介して相手に音声で送

信する。相手のマルチメディア公衆電話25では、直接の音声又は合成した音声を受信する。

【0034】したがって、図2の構成により、スクリーンから文字入力することにより、これを文章理解と音声合成により音声に変換し、これを音声信号として相手側に送ることにより、言語障害があっても会話が難しい難聴者の場合でもメモ機能などで話の内容を伝えるために文字で書き表した内容を音声で相手に伝える事ができる。

【0035】図7は、文字／音声変換スイッチの設置例である。このスイッチは、字幕を表示させるための字幕ボタンと、文字を入力して音声に変換させるための文字入力ボタンである。図8は、画面レイアウト例である。この画面はマルチメディア機能の他、映像部31、音声を字幕に切り替えた受信内容テキスト画面32、相手に文字を入力して音声で内容を送信するための送信内容テキスト画面33、入力した内容を送信する時、音声を男性の声か女性の声かを選択する音声モード34の4つからなる。

【0036】なお、本発明は、図1の構成又は図2の構成の両方の装置を搭載するマルチメディア公衆電話システムとする他に、図1又は図2の一方の装置のみを単独に搭載するマルチメディア公衆電話システムとすることもできる。

【0037】図9は、相手の音声を字幕表示で受信する機能と、文字入力を相手に音声で送信する機能の両方を持つマルチメディア公衆電話システムのインタフェース構成例を示す。

【0038】同図の音声入力部は、図10に処理手順で示すように、マイクロフォンから音声を入力し（S1）、ノイズフィルタで音声に交じっている雑音を除去し（S2）、音声入力ボードによりクリアな音声をデジタル信号に変換し（S3）、音声入力ドライバの制御の基にニューラルネットワークによる学習機能を有して音声認識し（S4）、ファジィ推論部の推論により認識された音声信号に冗長な信号を除去し（S5）、知識ベースと知識ベースファイルと自動プログラミングエディタにより認識された音声信号が知識ベースファイルに無い場合に学習処理を行い記憶し（S6）、ウィンドウコントロールにより最終的に音声信号処理が完了・認識された結果をウィンドウのエディタアプリケーション画面に表示する（S7）。

【0039】図9の文字入力部は、図11に処理手順で示すように、テキストエディタとペンにより文字入力し（S11）、入力された文字のパターン処理で認識するため、文字パターン認識部と学習部と専用知識ベース及び汎用知識ベースエディタにより、入力した文字のパターンやデータベース（db）に記憶されていない単語を学習して記憶し、そして文字のパターンが判別されることと認識する（S12）。次いで、認識された文字（デジタル）を音声に変換する（アナログ）ため、音声台

成システムにより男性の声又は女性の声で音声合成し（S13）、音声合成処理を完了した結果を音声出力部（マイクロホン）で出力する（S14）。

【0040】

【発明の効果】以上のとおり、本発明によれば、従来のマルチメディア公衆電話機能に音声理解システムを追加し、相手の音声を字幕に変換処理してディスプレイ上に表示させ、さらにメモ機能に音声理解機能を追加することでメモに記述した文字を音声に変換して通話者に伝達できるようにしたため、以下の効果がある。

【0041】（1）マルチメディア機能の利用範囲を広げる事が可能になる。

【0042】（2）会話中、重要と思われる点を字幕で表示させる事ができるため、聞き違いなどのリスクを回避する事が可能になる。

【0043】（3）電話などの会話と共に様々な情報範囲を広げる事が可能になる。

【0044】（4）聴覚障害者が、字幕や音声変換機能によって電話での会話が可能となる。

【0045】（5）文字を入力して音声に変換して相手へ送信する時、男性の声か女性の声かを選択する事ができる。

【図面の簡単な説明】

【図1】本発明の実施形態を示すマルチメディア公衆電話システムのブロック構成図（その1）。

【図2】本発明の実施形態を示すマルチメディア公衆電話システムのブロック構成図（その2）。

【図3】図1における処理フロー（その1）。

【図4】図1における処理フロー（その2）。

【図5】図2における処理フロー（その3）。

【図6】図2における処理フロー（その4）。

【図7】実施形態におけるスイッチ設置例。

【図8】実施形態における画面レイアウト例。

【図9】実施形態におけるインタフェース構成例。

【図10】図9における音声入力部の処理手順。

【図11】図9における文字入力部の処理手順。

【図12】現在のマルチメディア公衆電話の図。

【符号の説明】

11…音声入力装置

12…雑音除去ノイズフィルタ

13、24…通信システム

14…音声受信装置

15…音声／文字変換スイッチ

16…音声認識装置

17…音声認識生成装置

18…ディスプレイ表示装置

19…文字／音声変換スイッチ

20…文字入力装置

21…文章理解装置

22…音声合成装置

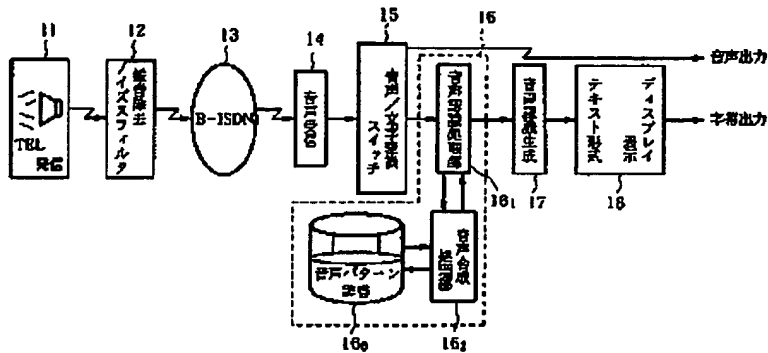
23…音声変換装置

* * 25…マルチメディア公衆電話

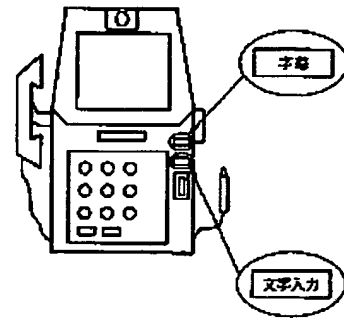
【図1】

【図7】

マルチメディア公衆電話システムのブロック図(その1)

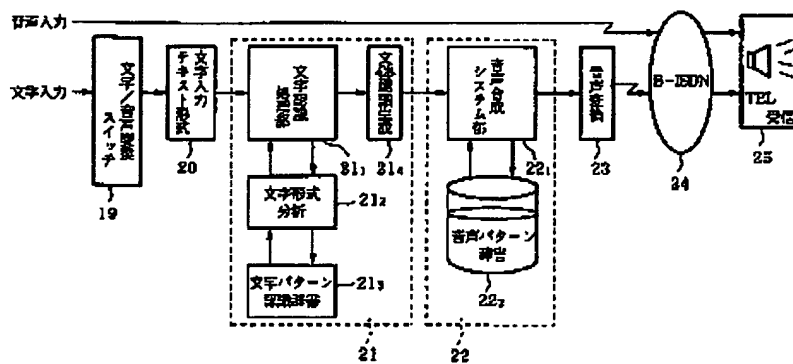


スイッチ設置例



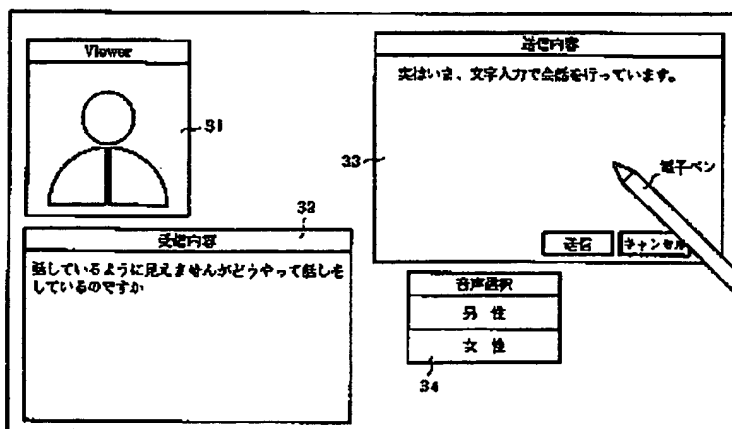
【図2】

マルチメディア公衆電話システムのブロック図(その2)



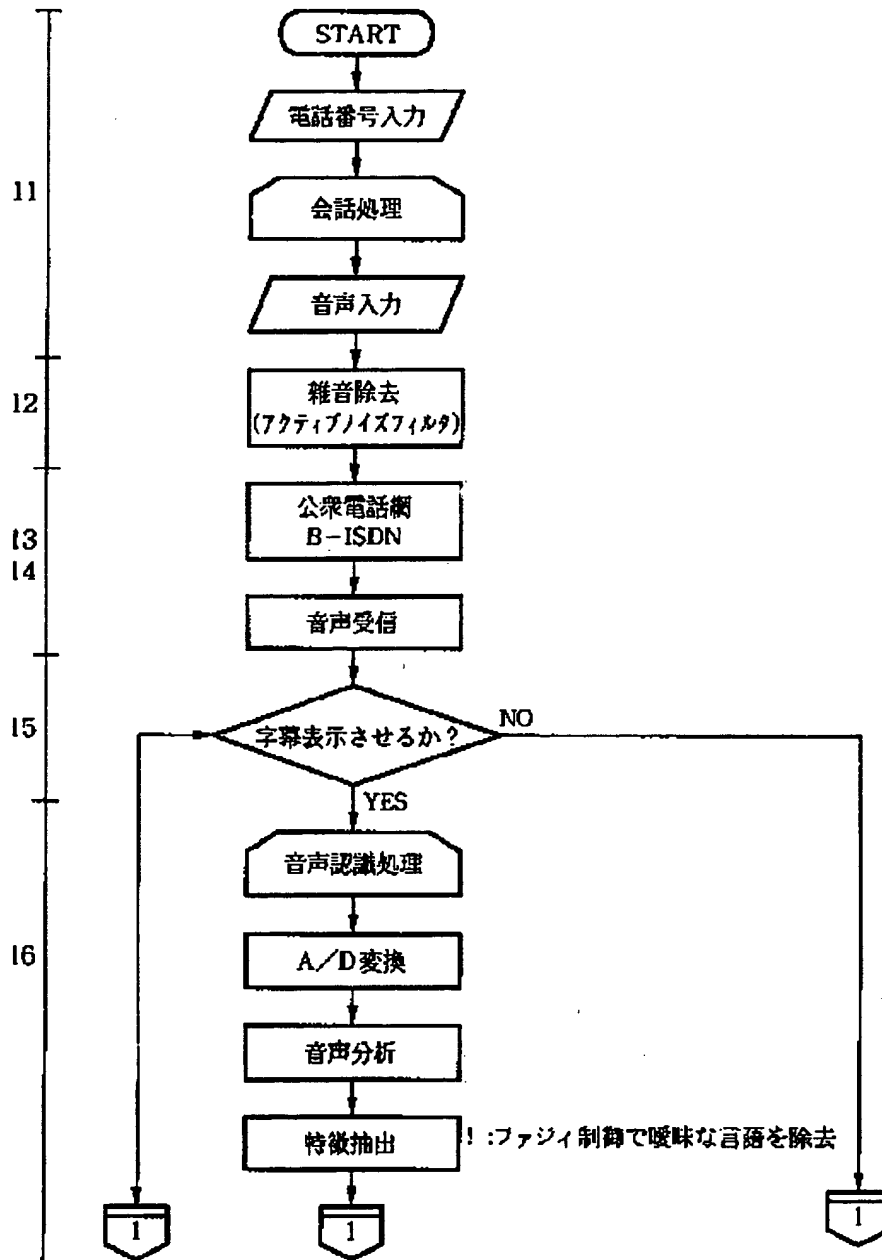
【図8】

画面レイアウト図



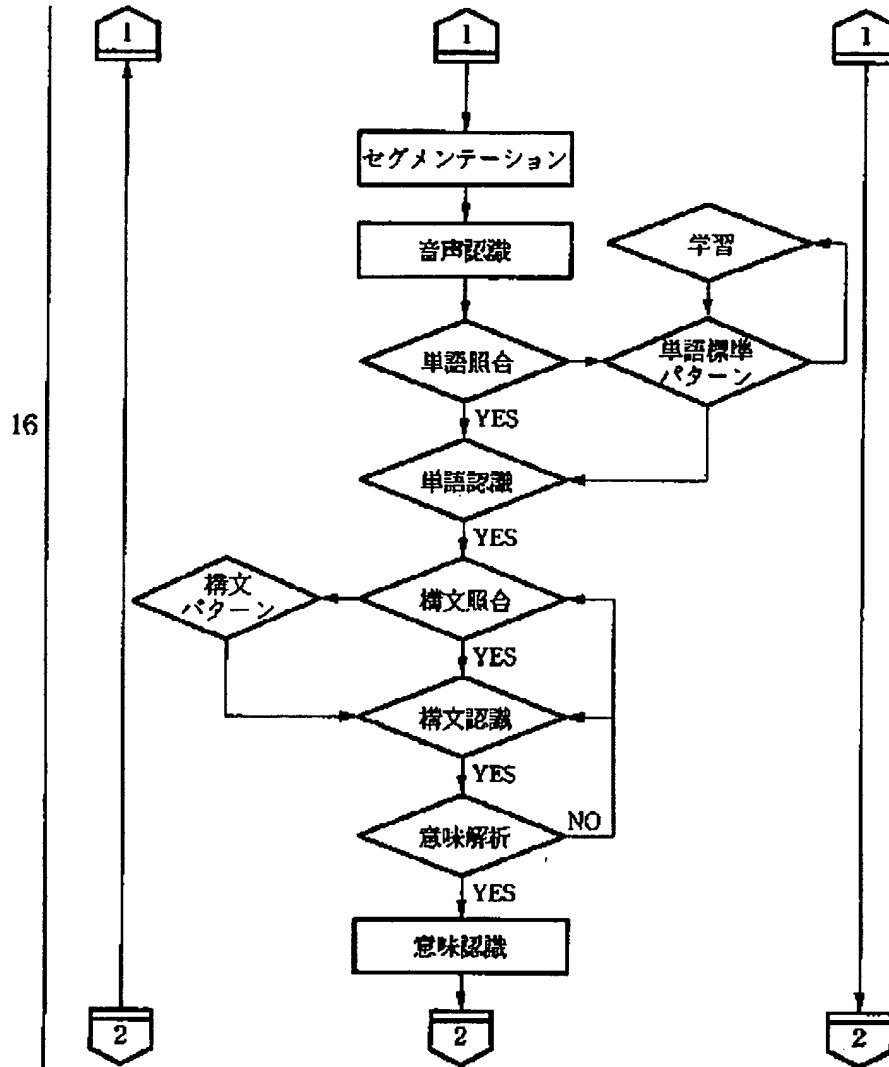
【図3】

処理フロー(その1)

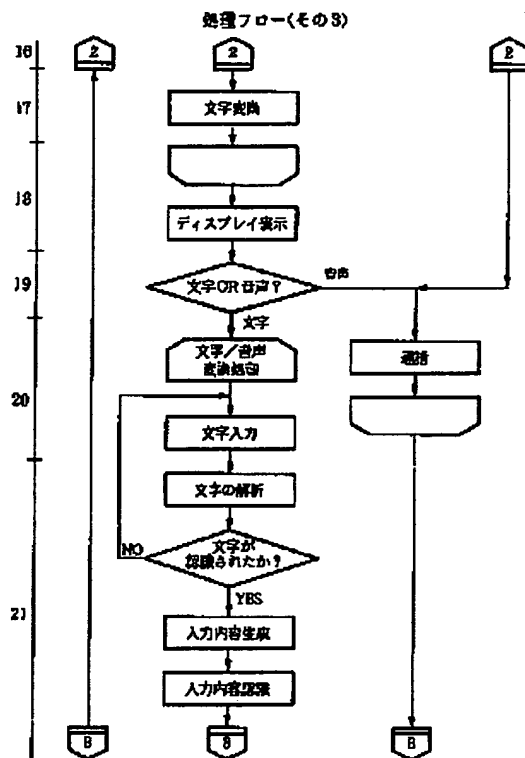


【図4】

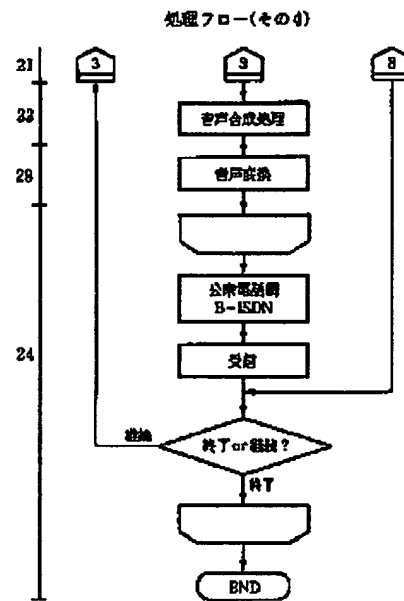
処理フロー(その2)



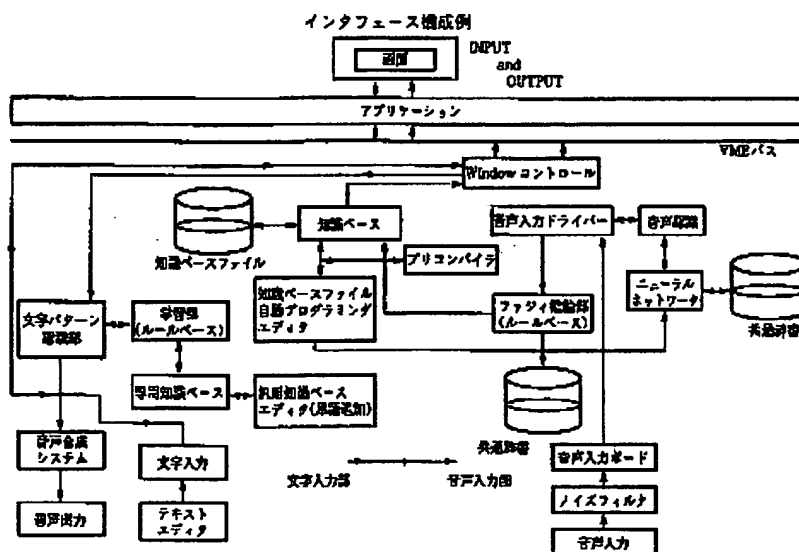
【図5】



【図6】

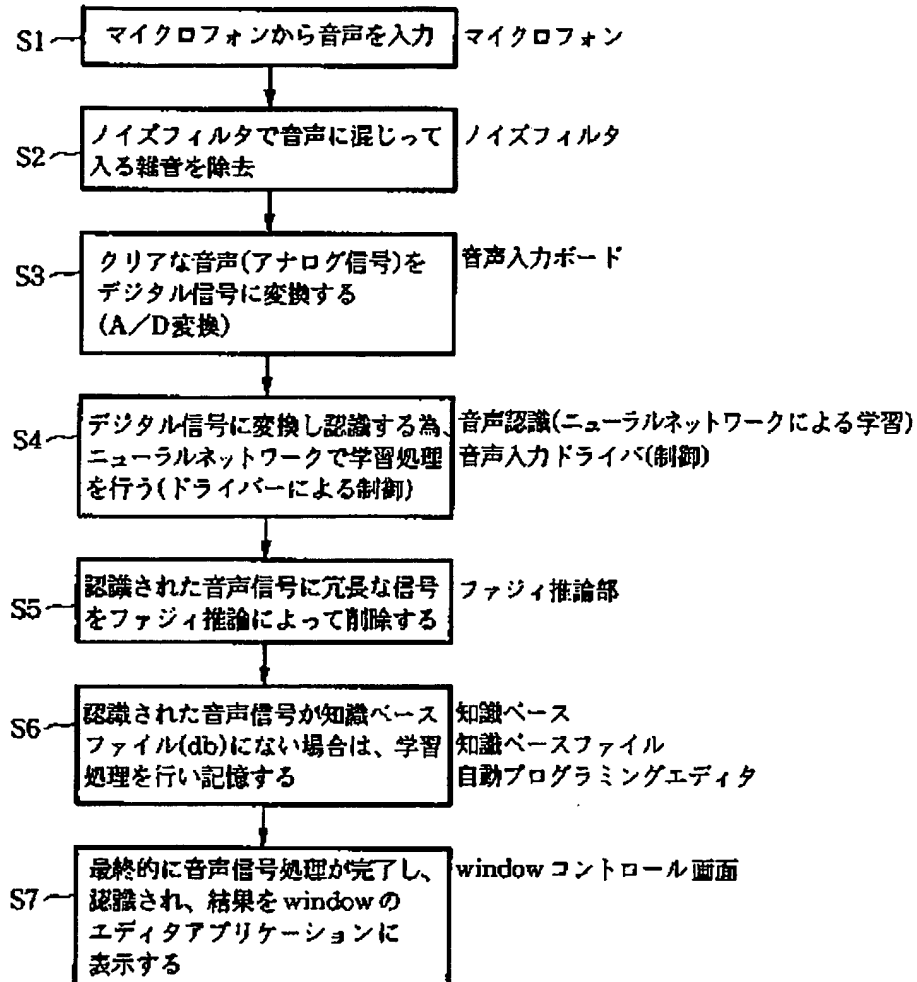


【図9】



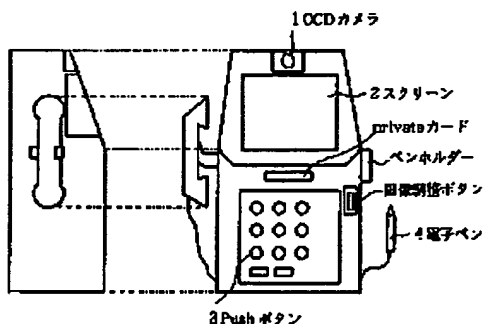
【図10】

音声入力部の処理手順



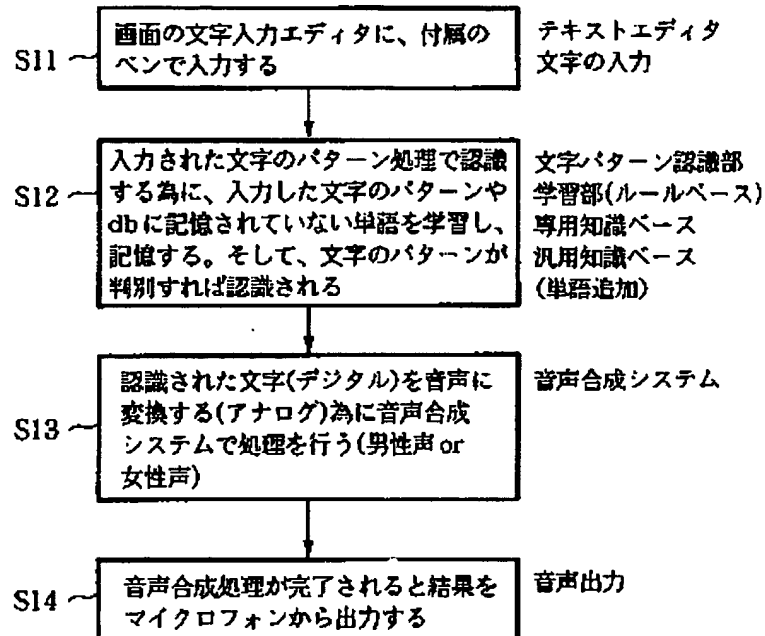
【図12】

現在マルチメディア公衆電話の図



【図11】

文字入力部の処理手順



CLAIMS

[Claim(s)]

[Claim 1] The multimedia public telephone system which has the voice and the I/O means of an image which are characterized by providing the following, and means of communications, and enabled the telephone call using voice and the image mutually. A sound signal receiving means to receive a telephone call person's voice through communication system. The voice / transliteration switch which chooses whether the aforementioned sound signal is outputted in an output or a title by voice as it is. The voice recognition unit which recognizes the sound signal which received when the aforementioned switch chooses a title output. Speech recognition generation equipment which changes into an alphabetic data the voice which the aforementioned voice recognition unit has recognized, and is displayed in a title on the aforementioned display.

[Claim 2] The multimedia public telephone system which has the voice and the I/O means of an image which are characterized by providing the following, and means of communications, and enabled the telephone call using voice and the image mutually. The character / voice conversion switch which chooses whether a telephone call person's voice is transmitted to a telephone call partner as it is, or it transmits in a character. Text understanding equipment which carries out recognition generation of the character string which the telephone call person inputted as a style by character recognition when the aforementioned switch chooses character transmission. The voice synthesizer which changes into a voice wave the text which the aforementioned text understanding equipment generated by speech synthesis. The voice inverter which changes into a sound signal the voice wave which the aforementioned voice synthesizer compounded, and transmits to a telephone call partner through communication system.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[The technical field to which invention belongs] this invention relates to the multimedia public telephone system which enabled it to transmit and receive an image etc. other than voice in a public telephone system.

[0002]

[Description of the Prior Art] Functions, such as having that conversation is made simultaneously, looking at a partner's face, having a FAX function, having the circumference map and the information guidance function of an installation, and a memorandum column function with the image I/O device which becomes an image input by CCD camera 1 and an image output by the screen 2, as the present multimedia public telephone system is shown in drawing 12, are added, and various kinds of functions, such as that a weather report can hear further and ordering the information on each event, are. As for the operation input of this public telephone, a push button 3 and a pointing device (a pen input and touch screen method) 4 are used.

[0003]

[Problem(s) to be Solved by the Invention] The following technical problems occur in the conventional multimedia public telephone system.

[0004] (1) In the case of a hypacusis person, conversation cannot be heard even if it can see a telephone call person's face.

[0005] (2) A telephone call person's voice (the content of conversation) cannot be read as a character.

[0006] (3) When conversation is a difficult hypacusis person, in order for there to be mogilalia, and to tell the content of the talk by the memorandum function etc., the content expressed in writing in written form cannot be told to a partner with voice.

[0007] (4) A healthy person has not heard wrong or caught [all busy and], either, and may interrupt the talk.

[0008] As mentioned above, even if the conventional system has various multimedia functions, the range which can be used for a hearing-impaired person is restricted.

[0009] The purpose of this invention is to provide a hearing-impaired person etc. with the multimedia public telephone system which improved user-friendliness.

[0010]

[Means for Solving the Problem] In order that this invention may solve the above-mentioned technical problem, it adds the SUS to the conventional function, carries out transform processing of a partner's voice to a title, and in order to make it display on a display and to also make the handicap of a language system ease further, it is the thing changes into voice the character described to the memorandum by adding a speech-understanding function to a memorandum function, and it enabled it to transmit to a telephone call person, and it carries out the following composition as the feature.

[0011] In the multimedia public telephone system which has voice, the I/O means of an image, and means of communications, and enabled the telephone call using voice and the image mutually (1st invention) When a sound signal receiving means to receive a telephone call person's voice through communication system, the voice / transliteration switch which chooses whether the aforementioned sound signal is outputted in an output or a title by voice as it is, and the aforementioned switch choose a title output, It is characterized by having the voice recognition unit which recognizes the sound signal which received, and speech recognition generation equipment which changes into an alphabetic data the voice which the aforementioned voice recognition unit has recognized, and is displayed in a title on the aforementioned display.

[0012] In the multimedia public telephone system which has voice, the I/O means of an image, and means of communications, and enabled the telephone call using voice and the image mutually (2nd invention) When the character / voice conversion switch which chooses whether a telephone call person's voice is transmitted to a telephone call partner as it is or it transmits in a character, and the aforementioned switch choose character transmission, The text understanding equipment which carries out recognition generation of the character string which the telephone call person inputted as a style by character recognition, It is characterized by having the voice synthesizer which changes into a voice wave the text which the aforementioned text understanding equipment generated by speech synthesis, and the voice inverter which changes into a sound signal the voice wave which

the aforementioned voice synthesizer compounded, and transmits to a telephone call partner through communication system.

[0013]

[Embodiments of the Invention] Drawing 1 and drawing 2 are the block diagrams of the multimedia public telephone system in which the operation gestalt of this invention is shown, and match and show the processing process of each part to drawing 3 - drawing 6 with the same sign. Hereafter, the composition and processing of each part are explained in detail.

[0014] (1) When a partner's voice is received by title display (drawing 1).

[0015] An audio input unit 11 is for inputting voice from the microphone of the earphone carried in a public telephone, and generates the signal into which the telephone call person inputted voice from the earphone. The normal-mode-rejection filter 12 removes the noise which is indistinguishable and enters into conversation by the noise filter.

[0016] By the ATM switching system which controls B-ISDN and the network of 2B+D with high speed and wide band nature for example, by the digital circuit, communication system 13 performs communications control and transmits the sound signal from an audio input unit 11 to the voice receiving set 14.

[0017] An addressee chooses whether voice / transliteration switch 15 is changed into whether the voice (analog signal) which received is received as voice as it is, and a character (digital signal), and is displayed on a display.

[0018] Speech-understanding equipment 16 recognizes and understands receiving voice, when selection operation of a switch 15 is made into a transliteration. In order that this speech-understanding equipment 16 may recognize and understand the voice which received, it is constituted by the system which performs speech recognition by the word and the syntactic and semantic analysis after the feature extraction which removes the ambiguous language by the voice analysis and the fuzzy control as technology which can respond to a speaker independence, and is equipped with the speech-recognition processing section 161, the speech-synthesis processing section 162, and a voice pattern dictionary 163.

[0019] With this equipment 16, as drawing 3 and the processing flow of drawing 4 show, after carrying out A/D-conversion processing of the inputted sound signal, frequency spectrum analyzes a voice input signal roughly with a voice analysis, the feature parameter of a sound signal is extracted and changed from a voice-analysis result by the feature extraction at time series, and segmentation processing performs the segmentation to a voice unit.

[0020] a feature extraction -- "-- like -, " and "***-", and well ["well"], it obtains, and fuzzy reasoning extracts a redundancy-word, it is removed, and when people and a man talk, only a really required word is extracted Furthermore, the daily required word is registered beforehand, and by the learning function, gradually, it recognizes and adds and it goes a behind needed word.

[0021] Subsequently, a phoneme sequence is acquired by performing word recognition by comparison with a voice standard pattern by speech recognition, and a word is recognized by collating with word collating and the word standard pattern which it has in the knowledge base about a phoneme sequence by word recognition. If a standard pattern does not exist in this recognition, study processing (recognition and registration) is performed

and word knowledge is added to the knowledge base.

[0022] Subsequently, if it analyzes whether there is any error in syntax by collating with the syntax pattern of the knowledge base by syntax collating and syntax recognition and there is an error about the recognized word, syntax recognition will be performed by re-verifying. Furthermore, it investigates whether it is semantically appropriate about the word syntax recognized by a semantic analysis and semantic recognition, and syntax-analysis processing and semantic-analysis processing are repeated until an appropriate result is obtained.

Characters convertible [with transliteration processing] including whether to be convertible for a character at these analyses are changed into the kanji or kana.

[0023] Returning to drawing 1 , speech recognition generation equipment 17 changes into the alphabetic data of text form the voice data the content of receiving voice has been recognized to be by speech-understanding equipment 16. A screen display of this alphabetic data is carried out to the screen 2 of a public telephone in a title with the display display 18.

[0024] Therefore, a telephone call person's conversation can be read as a character also by a hypacousis person's case by speech-understanding equipment's performing speech recognition for the voice which received, and displaying on a screen by making this into a character by composition of drawing 1 .

[0025] Moreover, risk aversion, like a healthy person also misses by title display during communication by telephone is made possible.

[0026] (2) When transmitting with voice against a character input (drawing 2).

[0027] A character / voice conversion switch 19 chooses a character button by the operation, when communicating from the content expressed in writing in written form by the memorandum function, in order that the addressee of conversation may answer a transmitting person.

[0028] If the character button of a switch 19 is pushed, the character input unit 20 will carry out a screen display of the memo pad (text form), and will enable the input of the content of the talk by the POINTITENGU device (electronic pen).

[0029] Text understanding equipment 21 analyzes the character string which carried out the handwriting input, recognizes it as a text, and is equipped with the style recognition processing section 211, the style formal analysis section 212, the character-pattern recognition dictionary 213, and the style recognition generation section 214.

[0030] With this text understanding equipment 21, if the character recognition processing section 211 analyzes a character from a character pattern and style form to a handwriting character input and the content of an input differs from the content of analysis as a processing flow is shown in drawing 5 , reinput will be received as correction, the content of an input of a handwriting character is recognized as a text, and the style recognition generation section 214 generates as a text.

[0031] A voice synthesizer 22 changes into a sound signal the text by which recognition generation was carried out with text understanding equipment 21 by speech synthesis, and is equipped with the speech-synthesis-system section 221 and the voice pattern dictionary 222.

[0032] A voice synthesizer 22 synthesizes voice with a rule composite system, it analyzes reading, a word, a clause boundary, etc. by the syntax analysis which referred to the

Japanese dictionary to the kanji kana mixture sentence, the semantic analysis, etc., determines each parameter of the duration of the accent of a phoneme sequence, intonation, and a phoneme with reference to a voice pattern dictionary from this analysis result, and synthesizes voice considering these as a parameter of a vocal tube model.

[0033] The voice transducer 23 changes a sound signal wave into voice. this conversion -- setting -- a telephone call person -- he chooses male voice or female voice on a system, and transmits with voice the voice which pushed and compounded the transmitting button to a partner through communication system 24 Direct voice or the compound voice is received in a partner's multimedia public telephone 25.

[0034] Therefore, by changing this into voice by text understanding and speech synthesis by carrying out a character input from a screen, and sending to the other party by making this into a sound signal by composition of drawing 2 , even when conversation is a difficult hypacusis person, in order for there to be mogilalia, and to tell the content of the talk by the memorandum function etc., the content expressed in writing in written form can be told to a partner with voice.

[0035] Drawing 7 is the example of installation of a character / voice conversion switch. This switch is a title button for displaying a title, and a character input button for inputting a character and making it change into voice. Drawing 8 is an example of a screen layout. This screen consists voice of four of the voice modes 34 in which male voice chooses female ****, when transmitting the transmitting content text screen 33 for inputting a character into the image section 31 besides a multimedia function, the receiving content text screen 32 which changed voice to the title, and a partner, and transmitting the content to them with voice, and the inputted content.

[0036] In addition, also let this invention be the multimedia public telephone system which it considers as the multimedia public telephone system carrying the equipment of both the composition of drawing 1 , or the composition of drawing 2 , and also carries independently only one equipment of drawing 1 or drawing 2 .

[0037] Drawing 9 shows the example of interface composition of a multimedia public telephone system with both the function to receive a partner's voice by title display, and the function transmitted with voice against a character input.

[0038] The voice input section of this drawing inputs voice into drawing 10 from a microphone, as procedure shows (S1). Remove the noise mixed with voice by the noise filter (S2), and clear voice is changed into a digital signal on a voice input board (S3). Speech recognition of the learning function by the NYURRU network is had and carried out to the basis of control of a voice input driver. (S4), A redundant signal is removed to the sound signal recognized by reasoning of the fuzzy reasoning section (S5). When there is no sound signal recognized with the knowledge base, the knowledge base file, and the automatic-programming editor in a knowledge base file, it memorizes by performing study processing (S6). The result sound signal processing has finally been completed and recognized to be by window control is displayed on the editor application screen of a window (S7).

[0039] As procedure shows to drawing 11 , in order to carry out a character input with a text editor and a pen (S11) and to recognize the character input section of drawing 9 by pattern processing of the inputted character, By the character-pattern recognition section,

the study section, the exclusive knowledge base, and general-purpose *****-SUEDITA, the word which is not memorized by the pattern or database (db) of the inputted character is learned and memorized, and it recognizes by the pattern of a character being distinguished (S12). Subsequently, in order to change the recognized character (digital) into voice (analog), it synthesizes voice in male voice or female voice by the speech synthesis system (S13), and the result which completed speech synthesis processing is outputted in the voice output section (microphone) (S14).

[0040]

[Effect of the Invention] Since add the SUS to the conventional multimedia public telephone function, and carry out transform processing of a partner's voice to a title, it is made to display on a display, the character described to the memorandum by adding a speech-understanding function to a memorandum function further is changed into voice and it enabled it to transmit to a telephone call person according to this invention the above passage, there are the following effects.

[0041] (1) It enables the use range of a multimedia function to extend.

[0042] (2) Since all busy and the point that it is thought that it is important can be displayed in a title, it becomes possible to avoid risks, such as a hearing it difference.

[0043] (3) It becomes possible to extend various information ranges with conversation, such as a telephone.

[0044] (4) The conversation of a hearing-impaired person is attained by telephone by the title or the voice conversion function.

[0045] (5) When inputting a character, changing into voice and transmitting to a partner, male voice or female **** can be chosen.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] The block block diagram of the multimedia public telephone system in which the operation gestalt of this invention is shown (the 1).

[Drawing 2] The block block diagram of the multimedia public telephone system in which the operation gestalt of this invention is shown (the 2).

[Drawing 3] The processing flow in drawing 1 (the 1).

[Drawing 4] The processing flow in drawing 1 (the 2).

[Drawing 5] The processing flow in drawing 2 (the 3).

[Drawing 6] The processing flow in drawing 2 (the 4).

[Drawing 7] The example of switch installation in an operation gestalt.

[Drawing 8] The example of a screen layout in an operation gestalt.

[Drawing 9] The example of interface composition in an operation gestalt.

[Drawing 10] Procedure of the voice input section in drawing 9 .

[Drawing 11] Procedure of the character input section in drawing 9 .

[Drawing 12] Drawing of the present multimedia public telephone.

[Description of Notations]

- 11 -- Audio input unit
- 12 -- Normal-mode-rejection noise filter
- 13 24 -- Communication system
- 14 -- Voice receiving set
- 15 -- Voice / transliteration switch
- 16 -- Voice recognition unit
- 17 -- Speech recognition generation equipment
- 18 -- Display display
- 19 -- A character / voice conversion switch
- 20 -- Character input unit
- 21 -- Text understanding equipment
- 22 -- Voice synthesizer
- 23 -- Voice inverter
- 25 -- Multimedia public telephone